BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents *will not* correct images, please do not report the images to the Image Problem Mailbox.

REMARKS/ARGUMENTS

By the present amendment, previous claims 10, 12 and 13 have been amended and new claim 14has been added. Support for amended claim 10 can be found in the application, for example on page 9, lines 9-11 and lines 21-23. Support for claim 14 can be found on page 10, line 23. The claims have also been amended in order to replace the term "OX-2" with "CD200" as the latter term has replaced the former term in the scientific literature. Support for the amendment can be found in the application, for example on page 2, lines 31-32. The amendments to the claims have been made without prejudice and without acquiescing to any of the Examiner's objections. The amendment does not contain new matter and its entry is respectfully requested.

The Official Action dated August 8, 2003 has been carefully considered. It is believed that the following comments represent a complete response to the Examiner's rejections and place the present application in condition for allowance. Reconsideration is respectfully requested.

35 USC §112, First Paragraph

The Examiner has objected to claims 10-12 under 35 USC §112, first paragraph as containing subject matter which was not described in the specification in such a way as to enable one skilled in the art to which it pertains, or with which it is most nearly connected, to use the invention. We respectfully disagree with the Examiner for the reasons that follow.

The present invention relates to the finding by the inventors that inhibiting the immune suppression caused by CD200 can be used to treat tumors. The data provided in the application demonstrates that increased expression of CD200 allows a tumor to evade the immune system which results in tumor growth and increased mortality (see Figures 22 and 23 and page 87, lines 14-16). On the other hand, administering antibodies to CD200 to animals having a tumor and increased CD200 levels results in improved tumor immunity, reduced tumor cell growth and improved survival (see Figure 24 and page 87, lines 17). This was demonstrated with two different types of tumor, an EL4

tumor which is a thymoma and a C1498 tumor which is a myeloid tumor. The claims have been amended to specify that the method of the invention is directed towards tumors that can be treated by inhibiting or preventing immune suppression caused by CD200.

The Examiner has questioned some of the data provided in the application. In particular, the Examiner notes that mice pre-immunized with CD80 transfected cells and then challenged with anti-OX2 antibody did not demonstrate any improvement in survival. This result is consistent with the amended claims and the explanation above as the CD80 transfected cells do not produce elevated CD200 (see Figure 25) and therefore antagonizing CD200 would not be effective in reducing tumor cell growth and enhancing survival. In contrast, the inventors have shown that administration of the CD86 transfected tumor cells results in elevated levels of CD200 (Figure 25) which suppresses the immune response and allows survival of the tumors. However, when such mice are treated with an antibody to CD200, the immune suppression mediated by CD200 is reduced, allowing the immune system to attack the tumor and prolong survival in the animals.

Therefore, we respectfully submit that the claims currently of record are supported in the application and would be sufficient to allow one of skill in the art to practice the invention. In particular, once a particular tumor has been identified for treatment, one of skill in the art could determine whether or not it can evade the immune system through the CD200 pathway. In particular, as is taught in the application, one can immunize an animal with the tumor and determine whether or not it increases CD200 expression as was done for the tumor cells used in the Examples. Further, one can also directly assess the tumor to see if it expresses CD200 using methods that are routine in the art. Again, the results presented in the application are clear and consistent in demonstrating that tumors that are associated with an increase in CD200 in the tumor-bearing animal can be treated with antibodies to CD200. Therefore, we disagree with the Examiner's statement that the results in the application "appear to be limited to a highly specialized set of circumstances".

We also disagree with the Examiner's statement that the treatment of diverse types of cancers is highly unpredictable. The data in the application also demonstrates that the method can be used to prevent tumors that form lung nodules which would be predictive of the treatment of a solid tumor. In addition, both EL4 and C1498 tumors are hematopoietic cell tumors and therefore would be predictive of all hematopoietic cell tumors and specifically leukemias. In this regard, claim 12 has been amended in order to specify hematopoietic cell cancers and new claim 14 specifies leukemias.

With respect to the Examiner's comments on the Examples on pages 91-92 and Figure 3, we submit that there is no difference between the group receiving blood alone and the group receiving vehicle and anti-CD200 antibody. These results are consistent with the present invention as administration of CD200 antibody would not lead to treatment of the tumor in that case as there is no elevation of CD200 in the animal.

The role of CD200 in tumor immunity appears consistent with the scientific literature. In this regard, we enclose an article by Rosenwald et al. (J. of Exp. Medicine, Vol. 194, No. 11, Dec. 3, 2001, 1639-1647) and an article by Bohen et al. (PNAS, Vol. 100, No. 4, Feb. 18, 2003, 1926-1930). In both these articles, it is demonstrated that CD200 is associated with the tumor phenotype. In particular, in Rosenwald et al., it is demonstrated in Figure 1 that chronic lymphocytic leukemia cells express MRC OX-2 which is synonymous with CD200. In Bohen et al., it is demonstrated in Table 4 that lymphomas that are resistant to treatment with rituximab have increased expression of MRC OX-2. In addition, one of the present inventors, Dr. Reginald Gorczynski, has demonstrated that EL4 tumors overexpress CD200.

The Examiner also comments that "it was well known at the time of the invention was made that eliciting protective immunity in animals prior to tumor challenge is not reasonably predictive that a therapy will also be therapeutic". We are not certain why the Examiner is commenting as such as in the present Examples, the animals are first immunized with the tumor prior to administration of the antibody (see page 84, lines 1-12). As a result, the data in the present application is a useful therapeutic model.

The therapeutic utility of the CD200 cascade has been well established by the inventors in several models including graft rejection, fetal loss, autoimmune disease as well as tumor immunity, which are all described in the present application as well as in issued U.S. patent nos. 6,338,851 and 6,652,858. Therefore, the inventors have provided sufficient data to evidence that modulating the CD2000 cascade has a credible and therapeutic utility and that they had possession of the invention at the time the application was filed.

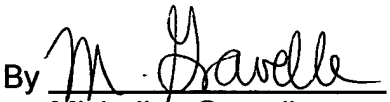
In view of the foregoing, we respectfully request that the objections to the claims under 35 USC §112, first paragraph as lacking written description, be withdrawn.

The Commissioner is hereby authorized to charge any deficiency in fees (including any claim fees) or credit any overpayment to our Deposit Account No. 02-2095.

In view of the foregoing, we submit that the application is in order for allowance and an early indication to that effect would be greatly appreciated. Should the Examiner like to discuss the matter, she is kindly requested to contact Micheline Gravelle at 416-957-1682 at her convenience.

Respectfully submitted,

BERESKIN & PARR

By 
Micheline Gravelle
Reg. No. 40,261

Bereskin & Parr
Box 401, 40 King Street West
Toronto, Ontario
Canada M5H 3Y2
Tel: 416-957-1682
Fax: 416-361-1398

Attachment

Relation of Gene Expression Phenotype to Immunoglobulin Mutation Genotype in B Cell Chronic Lymphocytic Leukemia

Andreas Rosenwald,[1] Ash A. Alizadeh,[2] George Widhopf,[5] Richard Simon,[6] R. Eric Davis,[1] Xin Yu,[1] Liming Yang,[1] Oxana K. Pickeral,[1] Laura Z. Rassenti,[5] John Powell,[7] David Botstein,[3] John C. Byrd,[8] Michael R. Grever,[9] Bruce D. Cheson,[10] Nicholas Chiorazzi,[11] Wyndham H. Wilson,[12] Thomas J. Kipps,[5] Patrick O. Brown,[2,4] and Louis M. Staudt[1]

[1]*Metabolism Branch, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, MD 20892*
[2]*Department of Biochemistry, the* [3]*Department of Genetics and* [4]*The Howard Hughes Medical Institute, Stanford University School of Medicine, Stanford, CA 94305*
[5]*University of California at San Diego, Department of Medicine, La Jolla, CA 92093*
[6]*Biometric Research Branch, Division of Cancer Treatment and Diagnosis, National Cancer Institute, National Institutes of Health, Bethesda, MD 20892*
[7]*Bioinformatics and Molecular Analysis Section, CBEL, CIT, National Institutes of Health, Bethesda, MD 20892*
[8]*Department of Medicine, Walter Reed Army Medical Center, Washington, D.C. 20307*
[9]*Department of Internal Medicine, Ohio State University, Columbus, OH 43214*
[10]*CTEP, Division of Cancer Treatment and Diagnosis, National Cancer Institute, National Institutes of Health, Bethesda, MD 20892*
[11]*North Shore-Long Island Jewish Research Institute, Manhasset, NY 11030*
[12]*Medicine Branch, Division of Clinical Sciences, National Cancer Institute, National Institutes of Health, Bethesda, MD 20892*

Abstract

The most common human leukemia is B cell chronic lymphocytic leukemia (CLL), a malignancy of mature B cells with a characteristic clinical presentation but a variable clinical course. The rearranged immunoglobulin (Ig) genes of CLL cells may be either germ-line in sequence or somatically mutated. Lack of Ig mutations defined a distinctly worse prognostic group of CLL patients raising the possibility that CLL comprises two distinct diseases. Using genomic-scale gene expression profiling, we show that CLL is characterized by a common gene expression "signature," irrespective of Ig mutational status, suggesting that CLL cases share a common mechanism of transformation and/or cell of origin. Nonetheless, the expression of hundreds of other genes correlated with the Ig mutational status, including many genes that are modulated in expression during mitogenic B cell receptor signaling. These genes were used to build a CLL subtype predictor that may help in the clinical classification of patients with this disease.

Key words: cDNA microarrays • gene expression profiling • leukemia • lymphocytic • chronic

Introduction

The observation that the rearranged Ig variable genes in chronic lymphocytic leukemia (CLL)* cells can either be unmutated or mutated suggested that CLL might comprise two different diseases that have been lumped together using

Address correspondence to Louis M. Staudt, Metabolism Branch, National Cancer Institute, Bldg. 10, Rm. 4N114, Bethesda, MD 20892. Phone: 301-402-1892; Fax: 301-496-9956; E-mail: lstaudt@box-l.nih.gov

*Abbreviations used in this paper: BCR, B cell receptor; CLL, chronic lymphocytic leukemia; DLBCL, diffuse large B cell lymphoma; PKC, protein kinase C.

standard diagnostic methods (1–3). Somatic hypermutation of Ig genes is a specialized diversification mechanism that is activated in B cells at the germinal center stage of differentiation (4, 5). Thus, it was suggested that CLL might include two disparate malignancies, one derived from an Ig-unmutated, pregerminal center B cell, and the other from an Ig-mutated B cell that has passed through the germinal center. This "two disease" model of CLL was further supported by the observation that Ig-unmutated and Ig-mutated CLL patients had distinctly different clinical courses (2, 3). This model predicts that Ig-unmutated and Ig-mutated CLL would not be highly related to each other in gene expression. A precedent for this model is found in the recent demonstration that another lymphoid malignancy, diffuse large B cell lymphoma (DLBCL), actually includes two distinct diseases that are morphologically indistinguishable but which have largely nonoverlapping gene expression profiles (6). An alternative hypothesis is that all cases of CLL have a common cellular origin and/or a common mechanism of malignant transformation. This model predicts that Ig-mutated and Ig-unmutated CLL cases should share a gene expression signature that is characteristic of CLL.

To test these two models, and to identify molecular differences between CLL patients that might influence their clinical course, we determined the gene expression phenotype of CLL on a genomic scale using Lymphochip cDNA microarrays (6, 7). Our data demonstrate that CLL, irrespective of the Ig mutational status, is defined by a characteristic gene expression signature, thus favoring the notion that all cases share some aspects of pathogenesis. Nonetheless, we found hundreds of genes differentially expressed between Ig-unmutated and Ig-mutated CLL providing the first molecular insight into the biological mechanisms that lead to the divergent clinical behaviors of these subgroups of CLL patients. The unexpected finding that B cell activation genes were differentially expressed between the two Ig-mutational subgroups in CLL suggests the intriguing possibility that signaling pathways downstream of the B cell receptor (BCR) contribute to the more aggressive clinical behavior of the Ig-unmutated subtype.

Materials and Methods

Microarray Procedures. Peripheral blood samples from CLL patients diagnosed according to National Cancer Institute guidelines (8) were obtained after informed consent and were treated anonymously during microarray analysis. 33 CLL patients studied had not received chemotherapy at the time of sample acquisition and four patients had received prior treatment. Ig mutational status was only studied in untreated patients. Leukemic cells from CLL blood samples were purified by magnetic selection for $CD19^+$ (Miltenyi Biotec) at 4°C before mRNA extraction and microarray analysis. Other mRNA samples from normal and malignant lymphoid populations have been described previously as have cell purification methods and array methods (6). All microarray experiments used the Cy5 dye to generate the experimental cDNA probe from mRNA of normal and malignant lymphocytes, and the Cy3 dye to generate the reference cDNA probe from mRNA pooled from nine lymphoma cell lines as described previously (6). Expression data presented in Figs. 1, 4, and 5 are available at http://llmpp.nih.gov/cll.

Initial microarray data selection was based on fluorescence signal intensity. Each selected data point either had 100 relative fluorescent units (RFU's) above background in both the Cy3 and Cy5 channels, or 500 RFU's above background in either channel alone. A supervised selection of genes preferentially expressed in CLL cells (see Fig. 1 A) was performed as follows. First, we used the fact that the majority of cell lines that were used to construct the reference pool of mRNA were derived from DLBCL. The percentage of CLL samples with expression ratio >3 relative to the reference cell line pool was calculated, and the same calculation was also performed for the DLBCL samples. Genes were selected for which >50% of the CLL samples, and <25% of the DLBCL samples, had ratios >3. Additionally, genes were selected if the average CLL ratio was greater than the average DLBCL ratio by greater than threefold. For Fig. 1 B, representative genes were chosen from Fig. 1 A by computing the average expression in CLL samples and the average expression in resting B cell samples (adult and cord blood B cells). CLL signature genes were chosen to be at least twofold more highly expressed in CLL than in resting B cells and CLL/resting B cell genes were chosen to be expressed equivalently (within twofold) in the two sample sets. Duplicate array elements representing the same genes were removed. Germinal center genes were chosen from a previous analysis (6).

RT-PCR. 500 ng poly-A^+ mRNA was used to generate first strand cDNA using Superscript (Life Technologies) together with random hexamers and oligo-dT primers. ZAP-70 oligonucleotide primers (5′ TCTCCAAAGCACTGGGTG 3′, 5′ AGCTGTGTGTGGAGACAACCAAG 3′) were then used for PCR amplification for 27 cycles.

Statistical Analysis. A two-group t-statistic on log2 expression ratios was used to measure the ability of each array element to discriminate between the two CLL mutational subtypes univariately. For multivariate subtype prediction, we used a linear combination of log2 expression ratios for array elements that were significant at the $P < 0.001$ significance level in the univariate analysis. The expression ratios were weighted in the linear combination by the univariate t statistics. The linear combination was computed for each sample and the average linear combination was computed for each CLL subtype. The midpoint of the two CLL subtype means was used as a cut-point for subtype prediction. For the cross-validation analysis, the subtype predictor was calculated by sequentially omitting one sample from the test set of cases, and using the remaining cases to generate the predictor. In Fig. 4 B, calculation of the P value from the permutation distribution of the t-statistic also demonstrated the high statistical significance of the differential gene expression between the CLL subtypes (data not shown). Classification was determined on all CLL cases with the exception of CLL-60 (Ig-unmutated) and CLL-21 and CLL-51 (minimally mutated cases).

In Fig. 5, the choice of B cell activation genes was made as follows. The B cell activation series of microarray experiments included several different stimulations with anti-IgM for 6, 24, and 48 h for each Lymphochip array element, we averaged the data at each activation time point, and then selected those elements that gave a twofold induction compared with the resting B cell average for at least one time point.

Results

The Gene Expression Signature of CLL. We profiled gene expression in CLL samples ($n = 37$) using Lymphochip

cDNA microarrays containing 17,856 human cDNAs (7). To facilitate comparison of each CLL mRNA sample with the others and with previously generated data sets, we compared gene expression in each CLL mRNA sample to a common reference mRNA pool prepared from lymphoid cell lines (6, 7). Using this strategy, the relative gene expression in the CLL cases could be compared with other B cell malignancies (DLBCL and follicular lymphoma) and of normal B cell and T cell subpopulations. Fig. 1 A presents expression data from 328 Lymphochip array elements representing ~247 genes that were selected in a supervised fashion (see Materials and Methods) to be more highly expressed in the majority of CLL samples than in DLBCL samples (n = 40). These genes fall into two broad categories, which are highlighted by representative genes in Fig. 1 B. Genes in the first category define a CLL gene expression "signature" that distinguishes CLL from various normal B cell subsets and from other B cell malignancies. The CLL signature genes were not expressed highly in resting blood B cells or in germinal center B cells. This group of genes includes several named genes not previously suspected to be expressed in CLL (e.g., Wnt3, titin, Ror1) as well as a number of novel genes from various normal and malignant B cell cDNA libraries. By contrast, CLL cells lacked expression of most genes that are preferentially expressed in germinal center B cells (Fig. 1 C). In addition to this set of CLL signature genes, CLL preferentially expressed a set of genes that distinguish resting, G0 stage blood B cells from mitogenically activated blood B cells and germinal center B cells that are traversing the cell cycle (Fig. 1 B). The expression of these resting B cell genes by CLL cells is consistent with the indolent, slowly proliferating character of this malignancy.

One of these resting B cell samples was prepared from human cord blood that is enriched for B cells bearing the CD5 surface marker, a B cell subpopulation that has been proposed to be the normal counterpart of CLL. The cord blood B cells were >80% $CD5^+$ by FACS® analysis (data not shown) whereas resting B cells from adult blood are 10–20% $CD5^+$ (9). We did not observe notably higher expression of the CLL signature genes in the cord blood B cell sample than in the adult B cell sample (Fig. 1) and no overall correlation in the expression of genes in Fig. 1 was observed between CLL and either adult or cord blood B cells (Pearson correlation coefficients −0.27 and −0.21, respectively). Thus, our gene expression profiling analysis does not provide support for the hypothesis that the $CD5^+$ B cell is a CLL precursor. It is certainly possible, however, that the expression of the CLL signature genes might be due to the oncogenic mechanisms of CLL and therefore might not be a feature of any normal B cell subpopulation.

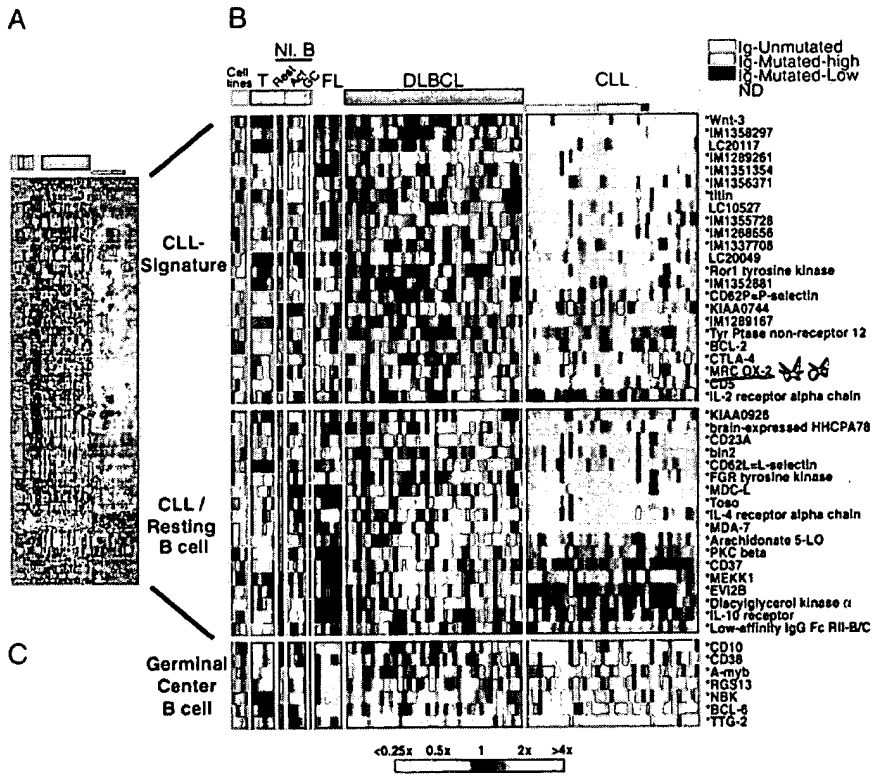


Figure 1. Discovery of a common gene expression phenotype in CLL. (A) 328 Lymphochip array elements representing ~247 genes that were more highly expressed as mRNA in the majority of CLL samples relative to DLBCL samples. The expression data are presented as a matrix in which the rows represent individual genes, and the columns represent individual mRNA samples. The relative level of gene expression is depicted according to the color scale shown at the bottom. Gray squares indicate missing or excluded data. (B) Relative expression levels of selected genes from A in mRNA samples presented in the following order, from left to right: cell lines (JVM-HH, OCI-Ly10, OCI-Ly3, U937); T cells (adult, $CD4^+$, unstimulated; neonatal, $CD4^+$, unstimulated; fetal, $CD4^+$, unstimulated; adult, $CD4^+$, + PMA [P] and ionomycin [I]; neonatal cord blood T cells, + P and I; fetal, $CD4^+$, + P and I), resting B-cells (cord blood $CD19^+$ B cells; adult blood $CD19^+$ B cells), activated B cells (adult blood B cells, anti-IgM + CD40L 6 h; adult blood B cells, anti-IgM 24 h; adult blood B cells, anti-IgM + CD40L 24 h; adult blood B cells, anti-IgM + IL-4 24 h), adult blood memory B-cells ($CD27^+$), tonsil germinal center B cells, follicular lymphomas (n = 7), DLBCLs (n = 40), and CLL (n = 37). The CLL samples were grouped according to Ig mutational status as indicated. In the gene names, an asterisk denotes a sequence verified gene, IM indicates an IMAGE consortium (reference 29) clone identification number, and LC indicates an unsequenced Lymphochip clone identification number. (C) Relative expression levels of selected genes characteristic of the germinal center B cell differentiation stage.

Ig Mutational Status. The expressed Ig heavy chain genes were sequenced from 28 CLL cases and compared with known germ-line encoded Ig V_H segments as described previously (10) (Fig. 2 A). By convention, V_H sequences that matched known germ line sequences with >98% identity were considered unmutated, as any minor differences observed in this group were assumed to reflect genetic polymorphism (1–3). By this criterion, 16 CLL cases in our study set were unmutated. The remaining cases were further separated into a group of 10 highly mutated cases (<97% identity with any germ-line V_H segment) and a group of two cases that were minimally mutated (>97% but <98% identity with known germ-line V_H genes). CLL cases were grouped in Fig. 1 according to Ig mutational status as indicated. Although some variation in expression of the CLL signature and CLL/resting B cell genes was evident between CLL patients, most patients in each Ig mutational subtype highly expressed these genes at comparable levels. Furthermore, an unsupervised hierarchical clustering of the CLL cases using 10,249 Lymphochip array elements resulted in a clustering dendrogram in which the Ig-unmutated and Ig-mutated CLL cases were extensively intermingled (data not shown). Thus, the overall gene expression profiles of the two CLL subtypes were largely overlapping.

Segregation of our patients according to Ig mutational status revealed that Ig-unmutated CLL patients had a significantly worse clinical course, requiring earlier treatment, than the Ig-mutated CLL patients (Fig. 2 B), in keeping with previous reports (2, 3).

CLL Subtype Distinction Genes. Given the dramatically different clinical behavior of the Ig-unmutated and Ig-mutated CLL patients, it was evident that gene expression differences should be discernible between these groups. To both discover such genes and statistically validate their relationship to the Ig-mutational subgroups, we conducted the Ig mutational analysis independently and sequentially in two random subsets of our CLL patients (Fig. 3). The "training" set consisted of 10 Ig-unmutated cases and eight Ig-mutated cases. In this gene discovery phase, we assigned the minimally mutated CLL cases to the mutated class. The mean expression of each gene was then calculated for both mutational subgroups and the statistical significance of the difference of these means was determined. All genes that discriminated between the mutational subgroups at a significance of $P < 0.001$ ($n = 56$) were used to form a "predictor" that could be used to assign a CLL sample to a mutational subgroup based on gene expression (see Methods).

The performance of this CLL subtype predictor was initially tested using a cross-validation strategy (Fig. 3 A). One of the 18 CLL samples in the training set was omitted, the statistically significant genes were determined, and a predictor was calculated based on the remaining 17 samples. The omitted sample was then assigned to a CLL subtype based

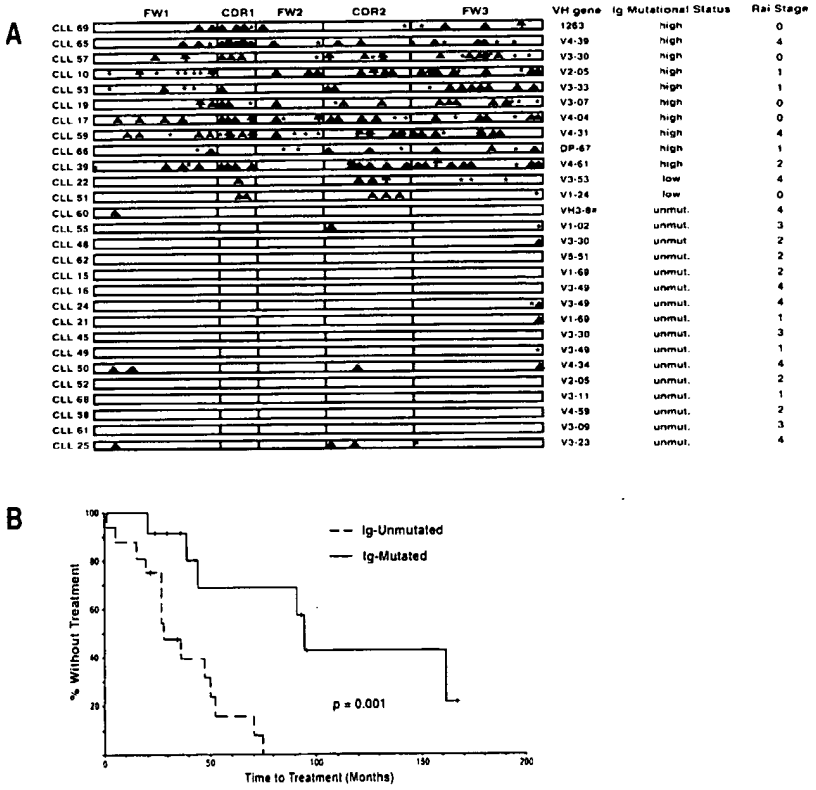


Figure 2. Analysis of somatic mutations in the Ig V_H genes in 28 CLL patients and correlation with their clinical courses. (A) V_H gene usage and distribution of replacement (▲) and silent (*) mutations in the complementarity determining regions (CDRs) and framework regions (FWs). #, V_H sequence most homologous to multiple cDNA sequences. (B) Kaplan-Meier curve comparing the time from diagnosis to treatment between CLL patients with mutated and unmutated V_H genes. Median time to treatment in Ig-mutated CLL: 95 mo; median time to treatment in Ig-unmutated CLL: 28 mo. The difference is significant at the $P = 0.001$ level (log-rank test).

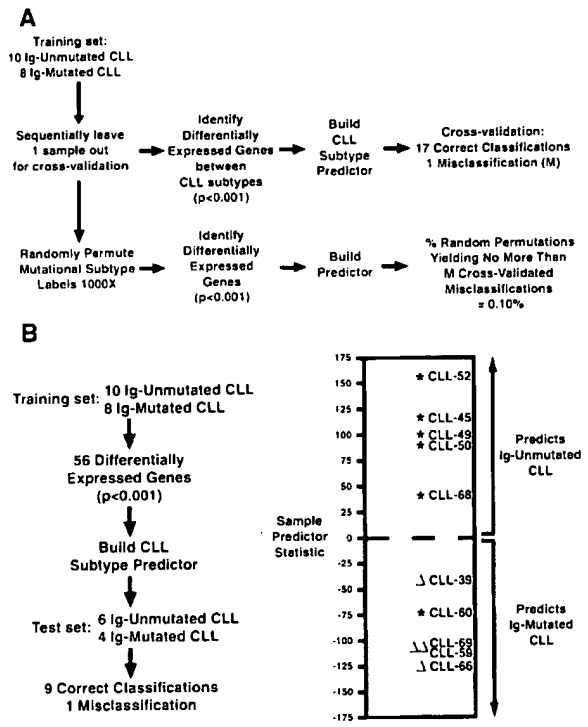


Figure 3. Statistical methodology for the creation and validation of an Ig-mutational status predictor in CLL. (A) Performance of the predictor using a cross-validation strategy. (B) Performance of the Ig-mutational subtype predictor in a test set of six unmutated (*) and four mutated CLL (Δ) samples.

on gene expression using this predictor. The Ig mutational status of 17 CLL samples was correctly assigned by this procedure with one misassignment. To test the statistical significance of this result, we created 1,000 random permutations of the assignments of CLL samples to the Ig mutation subgroups. For each permutation, the cross-validation process described above was repeated. Only one of the 1,000 random permutations generated a predictor that performed as well as the predictor based on the unpermutated data, demonstrating that the significance of the gene expression difference between the CLL subtypes was $P = 0.001$.

As a final test of the CLL subtype predictor, we determined the Ig mutational status of a "test" set of 10 additional CLL cases and used the predictor derived from the training set to assign the cases in this test set to a CLL subtype based on gene expression in a blinded fashion (Fig. 3 B). Nine out of ten of the test cases were correctly assigned, showing the ability of the CLL subtype predictor to correctly assign new CLL cases based on gene expression data that was not used to generate the predictor. The one misclassified CLL case (CLL-60) clearly was an outlier in gene expression (see below). Taken together with the cross-validation results, these data demonstrate that gene expression can define CLL subtypes that have different degrees of Ig mutation.

An important practical benefit of these findings would be to create a diagnostic test for the CLL subtypes based upon gene expression. In this regard, one of the most differentially expressed genes from the analysis of the training set of cases, ZAP-70, could classify all of the cases in both the training and the test set with 100% accuracy. Likewise, predictors based on two genes (ZAP-70 and IM1286077) or three genes (ZAP-70, IM1286077, activation-induced C-type lectin) discovered using the training set formed CLL subtype predictors that performed with 100% accuracy on the training set and test set of CLL cases.

We next expanded our search for CLL subtype distinction genes using data from both the training set and test set of CLL cases. The two CLL cases with minimal Ig mutations (CLL-22 and CLL-51) were excluded based on the possibility that their Ig sequences might actually represent as yet undescribed polymorphic V_H alleles. CLL-60 was excluded based on its unusual gene expression characteristics that led to its misclassification by the CLL subtype predictor. Fig. 4 A presents 205 Lymphochip array elements (~175 genes) that were differentially expressed between the CLL subtypes with a statistical significance of $P < 0.001$. Hierarchical clustering of the CLL cases based on expression of these genes placed the majority of Ig-unmutated CLL cases in one cluster and the Ig-highly mutated CLL cases in another. As expected, CLL-60 was more closely aligned with the Ig-mutated CLL cases, though it was an outlier from the major cluster of Ig-mutated CLL cases. Interestingly, both of the CLL cases with a low Ig mutational load were also outliers, though they were more closely related to the Ig-mutated CLL subtype than to the Ig-unmutated CLL subtype. These data define two predominant CLL subtypes that differ in the expression of hundreds of genes but also demonstrate that additional minor CLL subtypes may exist that have distinct gene expression profiles. Fig. 4 B highlights some of the genes that most strongly differentiate between the CLL subtypes. ZAP-70 was the most tightly discriminating gene, with an average 4.3-fold higher expression in Ig-unmutated CLL than in Ig-mutated CLL ($P < 10^{-6}$). RT-PCR analysis confirmed ZAP-70 expression in two Ig-unmutated CLL cases (CLL-48 and CLL-49), in contrast to CLL-66 and CLL-69 that were Ig-mutated (Fig. 4 C). Surprisingly, ZAP-70 expression was also observed in several B cell lines (LILA, LK-6, OCI-Ly2), but not in many others (Raji; Fig. 4 C, and data not shown).

Relationship between B Cell Activation and the CLL Subtype Distinction. Several of the CLL subtype distinction genes are known or suspected to be induced by protein kinase C (PKC) signaling, including activation-induced C-type lectin (11), MDS019, a very close paralogue of phorbolin 1 (12), and gravin, a scaffold protein that binds PKC and may regulate its activity (13). One mechanism by which PKC is activated in B cells is through BCR signaling (14). Therefore, we investigated whether the CLL subtype distinction genes are regulated during activation of blood B cells, using a gene expression database generated previously using Lymphochip microarrays (6). Strikingly, many of the

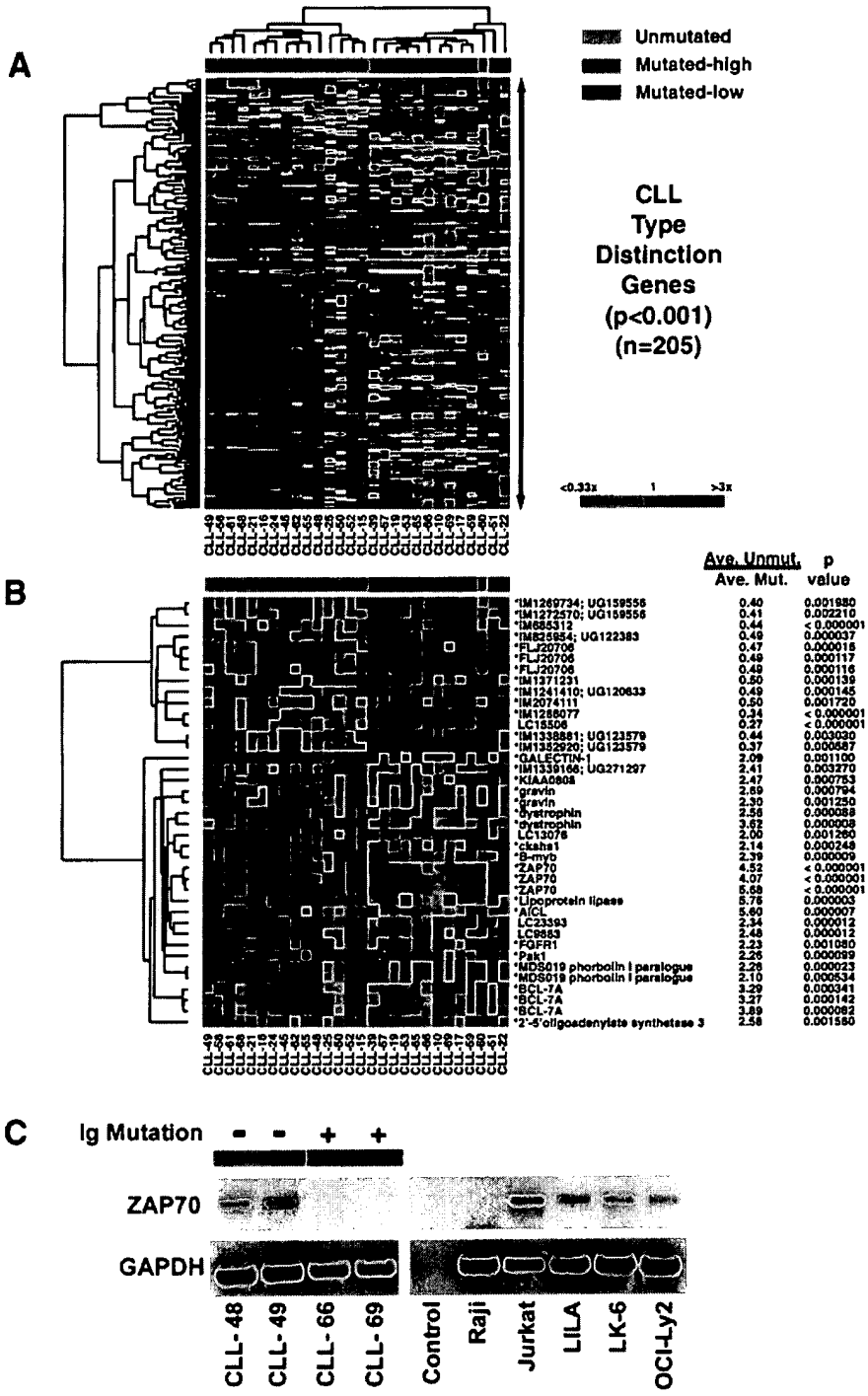


Figure 4. Relative gene expression levels of CLL subtype distinction genes. (A) Hierarchical clustering of gene expression data for 205 array elements representing ~175 genes that were differentially expressed between mutated and unmutated CLL samples ($P < 0.001$). (B) Hierarchical clustering of genes that most strongly discriminated between the CLL subtypes. Also shown for each gene is the ratio of mean expression of the gene in Ig-unmutated CLL samples (excluding CLL-60) versus mean expression in Ig-mutated (high) CLL samples, together with the P values (Student's t test) that quantitate the significance of the difference in mean expression between the two CLL subtypes. (C) RT-PCR analysis of ZAP-70 expression. Shown are data from two Ig-unmutated and two Ig-mutated CLL cases, a T cell line (Jurkat), various B cell lines found by microarray analysis to express ZAP-70 (LILA, LK6, OCI-Ly2), and a B cell line not expressing ZAP-70 (Raji; C, and data not shown). The control lane represents a reaction in which the reverse transcriptase was omitted.

genes that were more highly expressed in Ig-unmutated CLL were induced during activation of blood B cells (Fig. 5 A). Many of these genes encode proteins involved in cell cycle control (e.g., cyclin D2) or in cellular metabolism required for cell cycle progression (e.g., HPRT and other nucleotide modifying enzymes). Conversely, the majority of the genes that were expressed at lower levels in Ig-unmutated CLL were strongly downmodulated during B cell activation (Fig. 5 B). These results demonstrate that the CLL subtype distinction genes are enriched for genes that

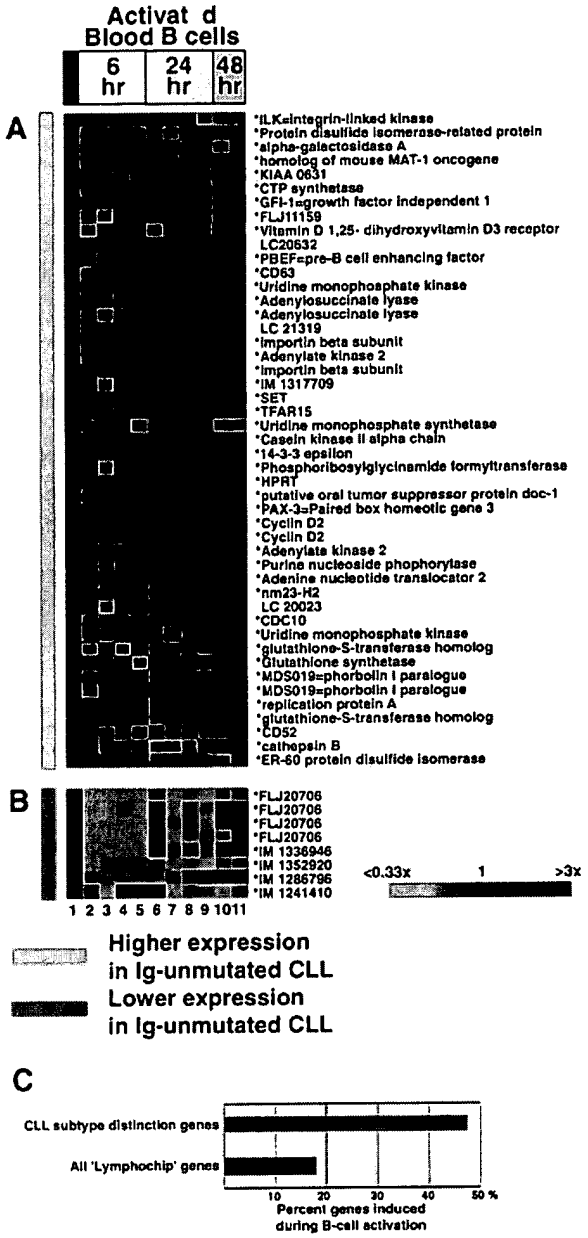


Figure 5. (A and B) Response of CLL subtype distinction genes during B cell activation. Gene expression data from the following B cell samples is depicted: 1, gene expression average from three resting B cell samples; 2, blood B cells, anti-IgM 6 h; 3, blood B cells, anti-IgM + CD40L 6 h; 4, blood B cells, anti-IgM + IL-4; 5, blood B cells, anti-IgM + CD40L + IL-4; 6, blood B cells, anti-IgM 24 h; 7, blood B cells, anti-IgM + CD40L 24 h; 8, blood B cells, anti-IgM + IL-4 24 h; 9, blood B cells, anti-IgM + CD40L + IL-4 24 h; 10–11, blood B cells, anti-IgM + CD40L + IL-4 48 h. (C) Percentage of CLL subtype distinction genes (red bar) and all 'Lymphochip' genes (blue bar) that are induced during B cell activation.

are modulated in expression by B cell activation. Indeed, 47% of the CLL subtype distinction genes were induced during B cell activation, whereas only 18% of all Lymphochip genes were in this category (Fig. 5 C).

Discussion

The comprehensive profiling of gene expression in CLL presented here provides a new molecular framework for understanding the etiology of this leukemia and the divergent clinical courses of these patients. Using genomic-scale gene expression profiling, we addressed a current controversy in CLL pathogenesis, namely whether this diagnosis comprises more than one disease entity. CLL patients have been subdivided based on the Ig mutational status of their leukemic cells (1–3), but it was unclear whether these patients had molecularly distinct diseases. Our data demonstrate that all CLL patients share a characteristic gene expression signature in their leukemic cells. These findings support a model in which all cases of CLL have a common cell of origin and/or a common mechanism of malignant transformation. In this model, the CLL-specific gene expression signature might represent the gene expression signature of a common normal precursor cell or it might reflect the downstream gene expression consequences of a common oncogenic event. These findings are in contrast to the previous observation that DLBCL consists of two disease entities that did not have overlapping gene expression outside of genes involved in proliferation and in the host response to the tumor (6).

Previously unsuspected features of CLL biology emerge from its gene expression profile, generating a wealth of hypotheses to guide future studies of this disease. CLL cells proliferate slowly in vivo, driven by unknown signals. Therefore, it is notable that Wnt-3 was highly, and selectively, expressed in CLL (Fig. 1 B). The Wnt gene family encodes secreted proteins that signal through cell surface receptors of the frizzled family to control development and mediate malignant transformation (15). Intriguingly, another CLL signature gene, Ror1, encodes a receptor tyrosine kinase with an extracellular domain that resembles a Wnt interaction domain of frizzled (16). Recently, Wnt-3 has been shown to promote proliferation of mouse bone marrow pro-B cells by initiating signaling events leading to transcriptional activation by LEF-1(17). Thus, CLL cells may use an autocrine mechanism of proliferation that is used normally by B cell progenitors.

We nevertheless also found that the expression of hundreds of other genes correlated with the Ig mutational status in CLL, providing insights into the biological mechanisms that lead to the divergent clinical behaviors of CLL patients. The most differentially expressed gene between the CLL subtypes was ZAP-70, a critical kinase that transduces signals from the T cell antigen receptor, and is preferentially expressed in normal T lymphocytes (18). Differential expression of ZAP-70 between CLL subtypes was therefore surprising since its expression in normal B cells has not been previously reported. However, by microarray analysis and RT-PCR analysis we found that ZAP-70 mRNA is highly expressed in some B lymphoma cell lines along with being differentially expressed by the CLL subtypes. A ZAP-70–related kinase, syk, transduces signals from the BCR (19), raising the possibility that ZAP-70

might alter BCR signaling in CLL cells. Another CLL subtype distinction gene, Pak1, could contribute to the resistance of CLL cells to apoptosis by phosphorylating Bad and thereby preventing Bad from inhibiting BCL-2 (20). FGFR1 is a receptor tyrosine kinase that can stimulate cellular proliferation after interaction with fibroblast growth factors. The higher expression of FGFR1 in Ig-unmutated CLL is intriguing given that CLL patients have elevated blood levels of basic fibroblast growth factor which can activate FGFR1 and block apoptosis in CLL (21, 22).

Intriguingly, CLL subtype distinction genes were enriched for genes that are modulated in expression during signaling of B cells through the BCR. One hypothesis raised by this observation is that the leukemic cells in Ig-unmutated CLL may have ongoing BCR signaling. Interestingly, the V_H repertoire usage in the Ig-unmutated and Ig-mutated CLL is distinct (1–3) and the combinations of V_H, D_H, and J_H gene segments rearranged in CLL cells are not random (1–3, 23, 24). These observations suggest that the surface Ig receptors of CLL cells may have specificity for unknown environmental or self-antigens. Indeed, CLL cells have been shown to frequently produce antibodies that bind classical autoantigens (25–27). The gene expression profiling data presented in this report raise the possibility that Ig-unmutated CLL cells may be continuously stimulated in vivo by antigen, giving rise to a gene expression profile that is reminiscent of BCR signaling. Indeed, CLL cells from patients with progressive disease were more readily stimulated by BCR cross-linking to synthesize DNA than were CLL cells from patients with stable disease (28). Although this study did not distinguish between Ig-unmutated and Ig-mutated CLL, the results are consistent with a differential ability of these subtypes to signal through the BCR. Alternatively, it is possible that Ig-unmutated CLL cells activate the same signaling pathways that are engaged during B cell activation as a result of genetic changes in the leukemic cells or by other pathological mechanisms.

An immediate clinical application of the present results would be in the differential molecular diagnosis of CLL. We demonstrated that as few as 1–3 genes could correctly assign patients to a CLL subtype with 100% accuracy. Thus, our results could be used to establish a quantitative RT-PCR test to diagnose the CLL subtypes and that would be easier to adopt clinically than DNA sequence analysis of Ig variable regions. Given the relatively benign course of Ig-mutated CLL, a simple diagnostic test based on gene expression would provide valuable prognostic information for CLL patients and could be used to guide treatment decisions.

Finally, our results suggest new therapeutic approaches to this currently incurable leukemia. First, the protein products of some of the CLL signature genes may present new targets for mAb therapy and for vaccine approaches to CLL. Second, the unexpected finding that B cell activation genes were upregulated in Ig-unmutated CLL patients suggests the intriguing possibility that signaling pathways downstream of the BCR may contribute to the more progressive clinical course of these patients. Thus, therapeutic targeting of these signaling pathways could specifically benefit those CLL patients that show gene expression evidence that these pathways are active.

We thank the Cancer Genome Anatomy Project (CGAP), led by Bob Strausberg and Rick Klausner, for help in constructing the Lymphochip microarray, and Christa Prange for providing CGAP cDNA clones. We also thank Rick Klausner for helpful discussions.

A. Rosenwald was supported by the Deutsche Krebshilfe, Bonn, Germany. Research at Stanford was supported by grants from the National Cancer Institute to D. Botstein and P.O. Brown, who is an Associate Investigator of the Howard Hughes Medical Institute. A. Alizadeh was initially supported by the Howard Hughes Medical Institute Research Scholar Program while at the National Institutes of Health and then by the Medical Scientist Training Program at Stanford University. This work was also supported by grants from the National Cancer Institute to T.J. Kipps and N. Chiorazzi (RO1CA 81554 and RO1CA 87956) and to the CLL Research Consortium.

Submitted: 1 August 2001
Revised: 23 August 2001
Accepted: 28 August 2001

References

1. Fais, F., F. Ghiotto, S. Hashimoto, B. Sellars, A. Valetto, S.L. Allen, P. Schulman, V.P. Vinciguerra, K. Rai, L.Z. Rassenti, et al. 1998. Chronic lymphocytic leukemia B cells express restricted sets of mutated and unmutated antigen receptors. *J. Clin. Invest.* 102:1515–1525.
2. Hamblin, T.J., Z. Davis, A. Gardiner, D.G. Oscier, and F.K. Stevenson. 1999. Unmutated Ig V(H) genes are associated with a more aggressive form of chronic lymphocytic leukemia. *Blood.* 94:1848–1854.
3. Damle, R.N., T. Wasil, F. Fais, F. Ghiotto, A. Valetto, S.L. Allen, A. Buchbinder, D. Budman, K. Dittmar, J. Kolitz, et al. 1999. Ig V gene mutation status and CD38 expression as novel prognostic indicators in chronic lymphocytic leukemia. *Blood.* 94:1840–1847.
4. Jacob, J., G. Kelsoe, K. Rajewsky, and U. Weiss. 1991. Intraclonal generation of antibody mutants in germinal centers. *Nature.* 354:389–392.
5. Klein, U., T. Goossens, M. Fischer, H. Kanzler, A. Braeuninger, K. Rajewsky, and R. Kuppers. 1998. Somatic hypermutation in normal and transformed human B cells. *Immunol. Rev.* 162:261–281.
6. Alizadeh, A.A., M.B. Eisen, R.E. Davis, C. Ma, I.S. Lossos, A. Rosenwald, J.C. Boldrick, H. Sabet, T. Tran, X. Yu, et al. 2000. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature.* 403:503–511.
7. Alizadeh, A., M. Eisen, R.E. Davis, C. Ma, H. Sabet, T. Tran, J. Powell, L. Yang, G. Marti, T. Moore, et al. 1999. The Lymphochip: a specialized cDNA microarray for the genomic-scale analysis of gene expression in normal and malignant lymphocytes. *Cold Spring Harbor Symp. Quant. Biol.* 64: 71–78.
8. Cheson, B.D., J.M. Bennett, M. Grever, N. Kay, M.J. Keating, S. O'Brien, and K.R. Rai. 1996. National Cancer Institute-sponsored working group guidelines for chronic lymphocytic leukemia: revised guidelines for diagnosis and treatment. *Blood.* 87:4990–4997.
9. Geiger, K.D., U. Klein, A. Brauninger, S. Berger, K. Leder,

K. Rajewsky, M.L. Hansmann, and R. Kuppers. 2000. CD5-positive B cells in healthy elderly humans are a polyclonal B cell population. *Eur. J. Immunol.* 30:2918–2923.
10. Bessudo, A., V. Cherepakhin, T.A. Johnson, L.Z. Rassenti, E. Feigal, and T.J. Kipps. 1996. Favored use of immunoglobulin V(H)4 Genes in AIDS-associated B-cell lymphoma. *Blood.* 88:252–260.
11. Hamann, J., K.T. Montgomery, S. Lau, R. Kucherlapati, and R.A. van Lier. 1997. AICL: a new activation-induced antigen encoded by the human NK gene complex. *Immunogenetics.* 45:295–300.
12. Madsen, P., S. Anant, H.H. Rasmussen, P. Gromov, H. Vorum, J.P. Dumanski, N. Tommerup, J.E. Collins, C.L. Wright, I. Dunham, et al. 1999. Psoriasis upregulated phorbolin-1 shares structural but not functional similarity to the mRNA-editing protein apobec-1. *J. Invest. Dermatol.* 113: 162–169.
13. Nauert, J.B., T.M. Klauck, L.K. Langeberg, and J.D. Scott. 1997. Gravin, an autoantigen recognized by serum from myasthenia gravis patients, is a kinase scaffold protein. *Curr. Biol.* 7:52–62.
14. Cambier, J.C., C.M. Pleiman, and M.R. Clark. 1994. Signal transduction by the B cell antigen receptor and its coreceptors. *Annu. Rev. Immunol.* 12:457–486.
15. Polakis, P. 2000. Wnt signaling and cancer. *Genes Dev.* 14: 1837–1851.
16. Saldanha, J., J. Singh, and D. Mahadevan. 1998. Identification of a Frizzled-like cysteine rich domain in the extracellular region of developmental receptor tyrosine kinases. *Protein Sci.* 7:1632–1635.
17. Reya, T., M. O'Riordan, R. Okamura, E. Devaney, K. Willert, R. Nusse, and R. Grosschedl. 2000. Wnt signaling regulates B lymphocyte proliferation through a LEF-1 dependent mechanism. *Immunity.* 13:15–24.
18. Chu, D.H., C.T. Morita, and A. Weiss. 1998. The Syk family of protein tyrosine kinases in T-cell activation and development. *Immunol. Rev.* 165:167–180.
19. Turner, M., E. Schweighoffer, F. Colucci, J.P. Di Santo, and V.L. Tybulewicz. 2000. Tyrosine kinase SYK: essential functions for immunoreceptor signalling. *Immunol. Today.* 21: 148–154.
20. Schurmann, A., A.F. Mooney, L.C. Sanders, M.A. Sells, H.G. Wang, J.C. Reed, and G.M. Bokoch. 2000. p21-activated kinase 1 phosphorylates the death agonist bad and protects cells from apoptosis. *Mol. Cell. Biol.* 20:453–461.
21. Aguayo, A., H. Kantarjian, T. Manshouri, C. Gidel, E. Estey, D. Thomas, C. Koller, Z. Estrov, S. O'Brien, M. Keating, et al. 2000. Angiogenesis in acute and chronic leukemias and myelodysplastic syndromes. *Blood.* 96:2240–2245.
22. Menzel, T., Z. Rahman, E. Calleja, K. White, E.L. Wilson, R. Wieder, and J. Gabrilove. 1996. Elevated intracellular level of basic fibroblast growth factor correlates with stage of chronic lymphocytic leukemia and is associated with resistance to fludarabine. *Blood.* 87:1056–1063.
23. Widhopf, G.F. II, and T.J. Kipps. 2001. Normal B cells express 51p1-encoded Ig heavy chains that are distinct from those expressed by chronic lymphocytic leukemia B cells. *J. Immunol.* 166:95–102.
24. Johnson, T.A., L.Z. Rassenti, and T.J. Kipps. 1997. Ig VH1 genes expressed in B cell chronic lymphocytic leukemia exhibit distinctive molecular features. *J. Immunol.* 158:235–246.
25. Borche, L., A. Lim, J.L. Binet, and G. Dighiero. 1990. Evidence that chronic lymphocytic leukemia B lymphocytes are frequently committed to production of natural autoantibodies. *Blood.* 76:562–569.
26. Sthoeger, Z.M., M. Wakai, D.B. Tse, V.P. Vinciguerra, S.L. Allen, D.R. Budman, S.M. Lichtman, P. Schulman, L.R. Weiselberg, and N. Chiorazzi. 1989. Production of autoantibodies by CD5-expressing B lymphocytes from patients with chronic lymphocytic leukemia. *J. Exp. Med.* 169:255–268.
27. Broker, B.M., A. Klajman, P. Youinou, J. Jouquan, C.P. Worman, J. Murphy, L. Mackenzie, R. Quartey-Papafio, M. Blaschek, P. Collins, et al. 1988. Chronic lymphocytic leukemic (CLL) cells secrete multispecific autoantibodies. *J. Autoimmun.* 1:469–481.
28. Aguilar-Santelises, M., H. Mellstedt, and M. Jondal. 1994. Leukemic cells from progressive B-CLL respond strongly to growth stimulation in vitro. *Leukemia.* 8:1146–1152.
29. Lennon, G., C. Auffray, M. Polymeropoulos, and M.B. Soares. 1996. The I.M.A.G.E. Consortium: an integrated molecular analysis of genomes and their expression. *Genomics.* 33:151–152.

Variation in gene expression patterns in follicular lymphoma and the response to rituximab

Sean P. Bohen*†‡, Olga G. Troyanskaya§, Orly Alter§, Roger Warnke¶, David Botstein§‖, Patrick O. Brown*‡, and Ronald Levy†

Departments of *Biochemistry, †Medicine, ¶Pathology, and §Genetics, and ‡Howard Hughes Medical Institute, Stanford University School of Medicine, Stanford, CA 94305

Contributed by David Botstein, December 24, 2002

Analysis of the patterns of gene expression in follicular lymphomas from 24 patients suggested that two groups of tumors might be distinguished. All patients, whose biopsies were obtained before any treatment, were treated with rituximab, a monoclonal antibody directed against the B cell antigen, CD20. Gene expression patterns in the tumors that subsequently failed to respond to rituximab appeared more similar to those of normal lymphoid tissues than to gene expression patterns of tumors from rituximab responders. These findings suggest the possibility that the response of follicular lymphoma to rituximab treatment may be predicted from the gene expression pattern of tumors.

Follicular non-Hodgkin's lymphoma (NHL) is an indolent B cell malignancy with an annual incidence exceeding 10,000 cases in the United States. Although follicular lymphoma (FL) is frequently responsive to treatment, therapy is very rarely, if ever, curative. Rituximab, a chimeric IgG1 monoclonal antibody directed at the B cell antigen CD20, has become a mainstay of treatment for low-grade NHL; >400,000 patients worldwide have been treated with rituximab. Phase II trials of rituximab in patients with refractory or relapsed low grade or follicular NHL demonstrated a 50% response rate (1).

Despite this extensive clinical experience, the mechanism of action of rituximab remains unclear, as does the nature of resistance (2). Among the proposed mechanisms are antibody-dependent cell-mediated cytotoxicity (3), complement-mediated cytotoxicity (4), and direct cytotoxicity through modulating CD20 function (5–7). The association with resistance to rituximab treatment of a low-affinity variant of the Fc receptor (8) is suggestive of an immune mechanism, and remains the only plausible hint about the nature of resistance.

In this study, we examined whether gene expression profiling using cDNA microarrays could reveal biological diversity among follicular lymphomas and, more specifically, whether gene expression patterns in tumors might predict sensitivity to rituximab treatment.

Materials and Methods

Patient Characteristics. Patients were included in this study based on the availability of freshly frozen lymph node biopsy material containing enough mRNA to allow cDNA microarray analysis. Only patients with samples that had been obtained before any systemic therapy were included. In all cases the pathological diagnosis was follicular non-Hodgkin's lymphoma [follicular small cleaved (grade 1), follicular mixed (grade 2), or follicular large cell (grade 3) histology]. Each patient received rituximab treatment with documentation of clinical outcome. In all cases, biopsy and pathology review were performed at Stanford University Medical Center. Rituximab treatment was administered either at Stanford University Medical Center or by an outside oncologist.

Microarray Procedures. Freshly frozen lymph node samples were obtained from patients who underwent excisional biopsy at Stanford University Medical Center between 1984 and 1997, who subsequently received rituximab between 1994 and 2000 and whose clinical response to rituximab treatment had been recorded. Tonsil and spleen samples were similarly obtained from patients treated at Stanford University Medical Center in 2000 or 2001. Biopsy samples were stored frozen in optimal cutting temperature compound. Poly(A)$^+$ mRNA was obtained from biopsy samples after homogenization of tissue with the FASTTRACK 2.0 kit (Invitrogen). An experimental cDNA probe incorporating Cy5 dye was generated from mRNA from malignant and normal lymphoid tissues; a common reference cDNA probe incorporating Cy3 dye was from mRNA derived from a panel of cell lines and probes were hybridized to cDNA microarrays as described (9, 10). Two types of microarrays were used. Some experiments in this study used Stanford Human arrays comprised of 38,431 DNA spots of 38,276 unique cDNA clones, representing ≈31,139 unique Unigene clusters of which 16,152 correspond to unique named genes. Some experiments were conducted with lymphochip (LC) microarrays comprised of 37,632 DNA spots with 32,876 unique cDNA clones, representing ≈17,622 Unigene clusters of which at least 10,250 are unique named genes. More detailed information regarding microarray methods, and data selection, and analysis, as well as searchable figures and microarray data files, can be found at http://genome-www.stanford.edu/rituximab.

Statistical Analysis. Before statistical analysis, individual data points were median centered for each cDNA clone and filtered for data quality and signal at least 2-fold above the median in two or three of the samples in each data set, as described in the web supplement. Agglomerative hierarchical cluster analysis was applied to the gene axis and to the sample axis as described (11). Hierarchical cluster analysis of LC data revealed a technical artifact that resulted in samples segregating by the date of the experiment. Further investigation revealed that this artifact was likely caused by differences in the calibration of the two scanners used to read the arrays. Singular value decomposition was used to remove the pattern corresponding to this artifact before analysis (12) after missing data were estimated by using a K-nearest neighbors (KNN) impute algorithm with 12 nearest values (13). Supervised analysis taking into account known outcome to rituximab treatment was performed by using Wilcoxon rank sum test to generate a rank list of genes whose corresponding mRNA levels differ significantly in rituximab responders versus nonresponders (14).

Results

Patient Characteristics. Tumor samples from 24 patients were analyzed in this study. No significant differences in age or treatment history were observed between responders and nonresponders (Table 1). All patients except one received at least one course of chemotherapy before receiving rituximab (range, 0–6 prior courses). One patient had received a shared anti-

Abbreviations: FL, follicular lymphoma; LC, lymphochip.

‖To whom correspondence should be addressed. E-mail: botstein@genome.stanford.edu.

www.pnas.org/cgi/doi/10.1073/pnas.0437875100

Table 1. Patient characteristics by response to rituximab treatment

Patient characteristics	PR/CR (n = 12)	MR/NR (n = 12)
Sex (M:F)	7:5	10:2
Age at diagnosis ± SD, years	45.3 ± 10.1	46.3 ± 13.2
Age at rituximab ± SD, years	52.2 ± 13.7	51.6 ± 9.2
Pathology		
FSC	7	9
FM	5	2
FLC	0	1
Mean courses of prior chemotherapy	2.4	2.4
Prior high dose chemotherapy	2	3

NR, no response; MR, minimal response; PR, partial response; CR, complete response; M, male; F, female; FSC, follicular small cleaved; FM, follicular mixed; FLC, follicular large cell. Age data are presented as mean age in years for each group ± standard deviation. None of the differences between the groups are statistically significant ($P > 0.35$ in all cases).

idiotype antibody as the sole prior treatment. For the treatment of FL, a course of rituximab typically consists of four weekly infusions of 375 mg/m^2; 21 of 24 patients received this dosing regimen, including all of the patients in the nonresponder group. In the partial response/complete response (PR/CR) group, two patients received a single dose of 375 mg/m^2 and one patient received eight weekly doses of 375 mg/m^2 with a documented near CR after 6 weeks of treatment. The overall response rate (CR+PR) for the patients in this study was 50%, which is similar to the overall response rate of 60% demonstrated for FL in the pivotal Phase II trial (1). Five patients achieved CR in response to rituximab, seven patients experienced PR, and 12 patients had no or minimal response.

Gene Expression Patterns Identify Two Subtypes of Follicular Lymphoma. An overview of the gene expression patterns from FL patients was generated by hierarchical cluster analysis of data from the first 16 FL patients analyzed together with samples from nonmalignant tonsil and spleen (Fig. 1*A*). The hierarchical cluster algorithm arranges tissue samples based on the degree of similarity in their gene expression patterns (11). In addition to tumor material obtained before treatment, we included tissue samples from later biopsies for four patients for whom such material was available. These repeat biopsies were obtained 2.5–6.5 years after the initial biopsy but before rituximab treatment. In all of the repeat biopsy cases, patients received one or more courses of chemotherapy between the initial and later biopsies and the histopathological diagnosis was the same for both biopsy samples.

To focus on genes differentially expressed in the samples, expression data from 2,037 unique genes whose expression differed 2-fold in at least three arrays were used for analysis. The dendrogram shown in Fig. 1*A* shows that the FL samples could be divided into two groups, with a subset of the FL samples exhibiting a gene expression pattern more similar to nonmalignant lymphoid tissues than to the other FL subtype. Hierarchical cluster analysis of the genes and tissue samples that contribute to the subclassification of FL samples are shown in Fig. 1*B*. The two subtypes of FL display gene expression patterns similar to or opposite those of nonmalignant spleen. It should be emphasized that the observed patterns of gene expression reflect all of the cells in the tumor, not only the tumor cells themselves. Interpretation of expression differences may include the differential presence or absence of other cell types, as has been found for breast cancer (9).

Rituximab Nonresponders Display Gene Expression Patterns Characteristic of Normal Lymphoid Tissue. Having divided the FL samples into two groups based on gene expression patterns, we sought to determine whether these subtypes correspond to clinical differences in individual patients, in particular, in response to rituximab treatment. We found that the rituximab nonresponders (orange in Fig. 1) were disproportionately distributed between the two FL clusters. Most of the rituximab nonresponders clustered with normal tonsil and spleen tissue ($P < 0.005$, Fisher's exact test). These findings indicate that there is biological diversity in FL lymph nodes from different patients before rituximab treatment and that the specific gene expression patterns defining the two groups described may be useful in predicting outcome to rituximab treatment.

Two features of the dendrogram in Fig. 1 indicate that the clustering of tissue samples based on gene expression patterns is reflective of biologically relevant similarities between samples. First, pairs of normal spleen and tonsil tissue cluster very closely together. Second, FL lymph node biopsies from the same individual cluster together in three of the four cases (see black bars in Fig. 1*A*) despite the passage of time and treatment with systemic chemotherapy between biopsies. An exception was the two samples obtained from patient 13 which did not cluster together (Fig. 1*A*, black arrows) with the pretreatment biopsy clustering with the responders' portion of the dendrogram and the later sample displaying a gene expression pattern more similar to that of rituximab nonresponders and normal lymphoid tissue. The latter sample was predictive of the patient's actual outcome after rituximab treatment; the patient did not respond. This observation suggests that the gene expression phenotype of FL can change over time.

Genes with Significantly Different Expression In Rituximab Responders Versus Nonresponders. To better understand differences in gene expression in involved lymph nodes from rituximab responders versus nonresponders, we used supervised statistical analysis to determine which genes had the most significant differences in expression between the two groups. Having observed that differences in gene expression in FL subtypes correlate with expression patterns in normal lymphoid tissues, samples from the original 16 patients and an additional 8 patient samples were analyzed on LC microarrays enriched for genes expressed in lymphoid cells and genes known or suspected to be important in the immune response or cancer (10). A list of genes whose expression differed between responders and nonresponders with a P value <0.005, as determined by Wilcoxon rank-sum test (14), was chosen. In cases where expression of a given gene was measured on both Stanford Human and LC microarrays, genes were included if the P value was less than 0.005 in either the Stanford Human or LC data set and a cDNA from the same Unigene cluster displayed a P value of <0.05 in the other dataset. Genes that were measured in only one data set were included if the P value is <0.005. The results of this analysis are presented in Tables 2 and 3. By these criteria, 71 genes had significantly higher expression in rituximab nonresponders versus responders; 53 named genes are on this list and 35 of the Unigene clusters represented were measured on both arrays. A total of 27 genes were more highly expressed by these criteria in rituximab responders; of these, 11 are named genes and 5 were measured on both arrays. Of note, CD20 expression was well measured and did not correlate with outcome to rituximab treatment.

Many of the genes with higher expression in tissue from rituximab nonresponders appear to be involved in cellular immune response and inflammation, specifically those encoding cytokine, tumor necrosis factor, and T cell receptor signaling, and complement proteins (Table 4). Given the limitations of available experimental evidence, this simplistic classification cannot account for the complexity of function or regulation of these genes; perhaps, mediators of cellular immune response, such T cells, macrophages, monocytes, and natural killer cells, may be relatively more abundant and/or more active in lymph node tissue of rituximab nonresponders. Several mRNAs for proteins involved in the complement cascade are more abundant

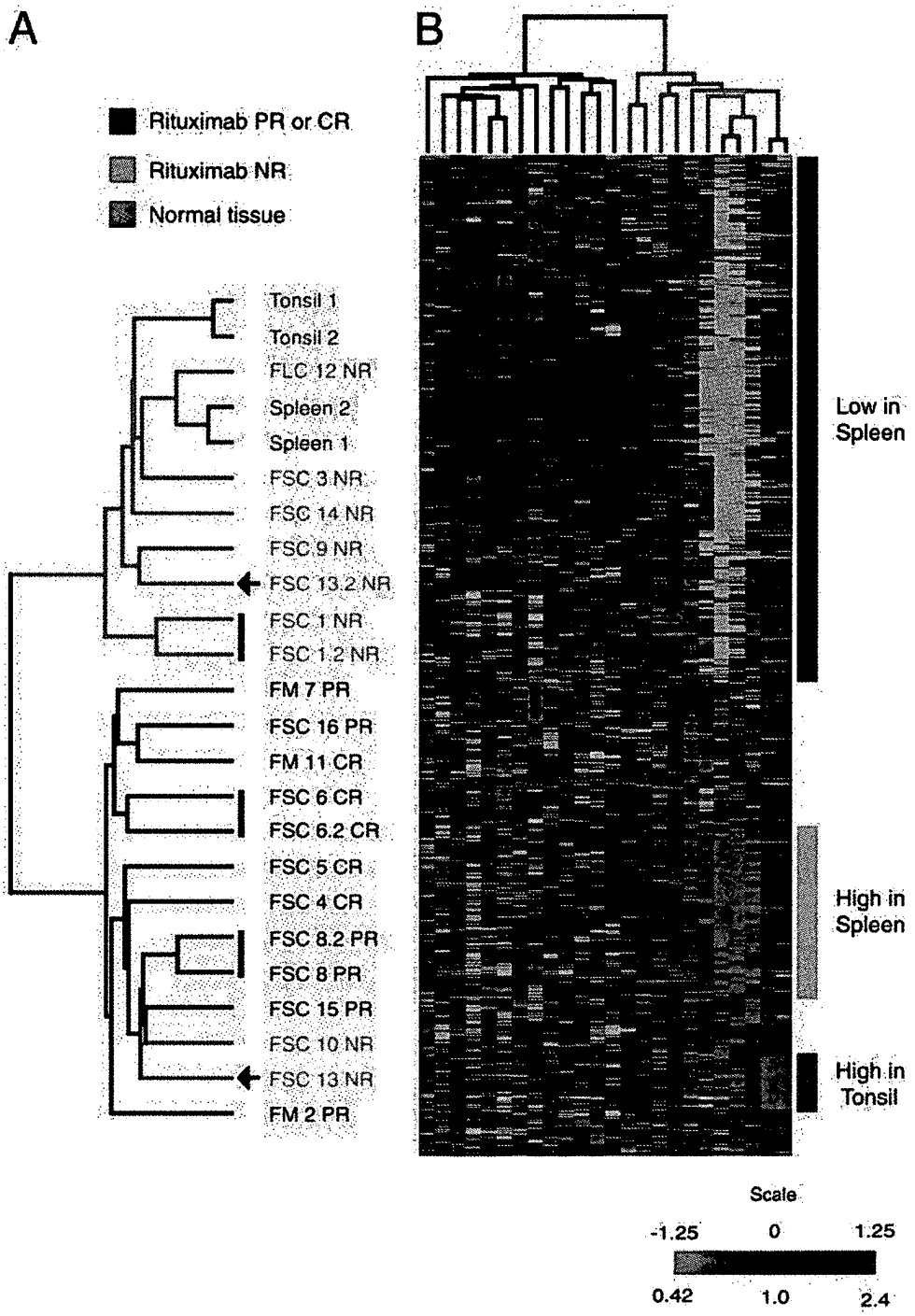


Fig. 1. Hierarchical cluster analysis of gene expression patterns in follicular lymphoma. (*A*) Patterns of gene expression in FL lymph nodes from rituximab responders cluster with normal lymphoid tissue. A total of 20 FL lymph node samples from 16 patients and four normal lymphoid tissues (two tonsil and two spleen) from four different patients were sorted by hierarchical clustering based on similarity of gene expression. The resulting dendrogram is shown. Patient pathological diagnosis or normal tissue type, and response to rituximab treatment are shown. Samples are color coded by response to rituximab (for FL samples) or normal tissue for simplicity. For four patients, paired samples are presented; biopsy samples obtained later are identified by patient number plus ".2". In three cases (patients 6, 8, and 18), paired samples clustered together (black bars). The two samples from patient 13 clustered on separate branches of the dendrogram (arrows). (*B*) Patterns of gene expression in FL samples and normal lymphoid tissue. Hierarchical clustering of genes and samples was performed. Variation of gene expression measured with 2,037 genes in 20 FL and four nonmalignant lymphoid tissue samples. Data are presented in a matrix format; each row represents a particular cDNA and each column is an individual FL or normal lymphoid tissue sample. For each sample, the ratio of the abundance of the mRNA measured by each cDNA clone to the median abundance of the mRNA across all tissue samples is represented by color in the corresponding cell in the matrix. Green represents transcript levels less than the median, black represents transcript levels equal to the median, and red represents transcript levels greater than the median. Color saturation represents the magnitude of the ratio relative to the median for each cell (see scale). Colored bars the to right of the matrix define groups of genes with similar expression patterns in normal lymphoid tissues, with blue denoting genes displaying relatively higher expression in spleen, purple denoting genes with higher expression in tonsil, and brown denoting genes with lower expression in spleen. Data can be viewed on our web site at http://genome-www.stanford.edu/rituximab/.

Table 2. Genes expressed at significantly higher levels in rituximab nonresponders versus responders

Gene or clone name	*P* value on LC	*P* value on SH
IMAGE:2021765	0.0002	N.M.
NPDC1	0.0002	0.0018
TNFRSF1B	0.0003	0.0150
CCND1	0.0248	0.0004
DUSP6	0.0047	0.0004
C4B	0.0006	0.0025
JUNB	0.0006	0.0041
PTPRM	0.0006	0.0035
FLJ20967	N.M.	0.0006
IMAGE:4100953	N.M.	0.0006
NGFRAP1	0.0007	0.0009
DAAM2	N.M.	0.0009
DKFZp564C2063	0.0248	0.0009
SSI-3	0.0010	0.0018
RET	0.0011	0.0320
MLPH	N.M.	0.0013
C1S	0.0013	0.0018
TRB@	0.0016	0.0404
CAV1	0.0065	0.0018
IFITM1	0.0047	0.0018
IMAGE:3899550	N.M.	0.0018
KIAA1223	N.M.	0.0018
RAB38	N.M.	0.0018
UACA	0.0040	0.0018
C1QR1	0.0019	0.0086
SNX9	0.0019	0.0114
IMAGE:1371537	0.0023	N.M.
MOX2	0.0023	N.M.
IMAGE:27277	N.M.	0.0025
FLJ23221	N.M.	0.0025
FLJ23705	N.M.	0.0025
LOC115908	N.M.	0.0025
LOC51087	N.M.	0.0025
NCF4	N.M.	0.0025
TYROBP	0.0077	0.0025
BHMT	0.0028	N.M.
PTD004	0.0028	0.0064
Similar to MM20	0.0028	N.M.
SLC21A9	0.0028	N.M.
SPP1	0.0028	N.M.
TNFSF10	0.0028	0.0195
CRY1	0.0033	N.M.
NK4	0.0033	0.0064

P values are determined by Wilcoxon rank sum test. SH and LC refer to different microarray configurations. "N.M." indicates that expression of this cDNA clone or Unigene cluster was not measured on the corresponding array type.

Table 3. Genes expressed at significantly higher levels in rituximab responders versus nonresponders

Gene or clone name	*P* value on LC	*P* value on SH
IMAGE:1582330	N.M.	0.0006
IMAGE:1475660	N.M.	0.0006
IMAGE:1504098	N.M.	0.0009
PRO0650	N.M.	0.0009
IMAGE:626773	N.M.	0.0009
IMAGE:460189	N.M.	0.0009
IMAGE:22374	N.M.	0.0013
Similar to MGR7	N.M.	0.0013
KIAA0317	N.M.	0.0018
H2BFG	N.M.	0.0018
RRM2B	N.M.	0.0018
DKFZp434K1210	N.M.	0.0018
IMAGE:1891596	N.M.	0.0018
BRI3BP	0.0470	0.0025
FREB	0.0470	0.0025
DKFZp564D113	N.M.	0.0025
UBQLN1	N.M.	0.0025
IMAGE:826372	N.M.	0.0025
DKFZp434B1620	N.M.	0.0025
GLE1L	N.M.	0.0030
H2BFB	0.0065	0.0035
ST14	0.0416	0.0035
IMAGE:752612	N.M.	0.0035
BLNK	0.0416	0.0048
IMAGE:1560875	N.M.	0.0048
BLCAP	N.M.	0.0048
IMAGE:814273	N.M.	0.0048

P values are determined by Wilcoxon rank sum test. SH and LC refer to different microarray configurations. "N.M." indicates that expression of this cDNA clone or Unigene cluster was not measured on the corresponding array type.

in nonresponders; the complement cascade can mediate direct cytotoxicity and/or act to stimulate cell-mediated cytotoxicity by opsonization of target cells. We were unable to identify a compelling physiological link among the 11 named genes with significantly increased expression in rituximab responders.

Discussion

Variation in gene expression patterns of FL lymph nodes suggests that it may be useful to divide these tumors into two groups. We suppose that these subtypes are reflective of basic biological differences between FL samples from different individuals. As mentioned before, the differences could be the result, at least in part, of the differential admixture of cells other than the malignant B cells that comprise most of the tumor. Thus, although it is quite possible that the subtypes identified simply represent molecular heterogeneity in the malignant cells, the finding that one subtype displays an expression pattern more similar to nonmalignant lymphoid tissue raises the possibility that the subtypes may, in fact, result from interactions between the malignant B cells and other cell types in the host. On the basis of the genes differing most in their expression between the two subtypes of tumors, it is possible that the differences may involve either the participation or the activities of cells involved in host immune response to the tumors. In any case, on the basis of this preliminary study, the subtype (and the response to rituximab) appears to be determined by the time of initial diagnosis and distinctive features of the gene expression pattern of a given individual's lymphoma appear to be recognizably retained over an interval of years.

Gene expression profiling studies of various malignancies have identified previously unappreciated subtypes within accepted pathological diagnoses and, in some cases, the prognoses of these subtypes varied significantly (10, 15–17). Several models of rituximab resistance are consistent with our findings. If rituximab nonresponders are capable of mounting a cellular immune response to their tumor before therapy, the FL cells may have been selected to evade this immune response and this may translate into an ability to evade any antibody-dependent cell-mediated cytotoxicity (ADCC) response subsequently triggered by rituximab. Alternatively, it has been proposed that FL cells may be able to recruit a cellular microenvironment similar to that in the germinal center, which facilitates the growth of the lymphoma cells (18). Following this line of reasoning, the list of genes with higher expression in rituximab nonresponders may indicate that these tumors may more effectively induce a growth stimulating microenvironment and this more favorable microenvironment may make these cells less sensitive to killing by

Table 4. A subset of genes with significantly higher expression in rituximab nonresponders versus responders are listed by biological function

Cytokine signaling
STAT4
STAT-induced STAT inhibitor 3
Secreted phosphoprotein 1 (osteopontin)
MRC OX-2 antigen
Small inducible cytokine, subfamily B, member 1 (GRO1)
Small inducible cytokine, subfamily A, member 2
Natural killer cell transcript 4
IFN-induced transmembrane protein 1 (LEU13)
IFN-induced transmembrane protein 2
IFN-induced transmembrane protein 3
IL-2 inducible T cell kinase
Protein tyrosine kinase binding protein
Neutrophil cytosolic factor 4
Complement
Complement component 4B
Complement component 1, s subcomponent
Complement component 1, q subcomponent, receptor 1
T cell receptor signaling
T cell receptor β
ζ-chain (TCR) associated protein kinase (70 kDa)
Tumor necrosis factor signaling
Tumor necrosis factor receptor superfamily, member 1B
Tumor necrosis factor, α-induced protein 2
Jun B protooncogene
Tumor necrosis factor (ligand) superfamily, member 10
Growth arrest and DNA damage-inducible gene, β
FOS protooncogene
p75NTR-associated cell death executor

A review of the literature for all 53 named genes in Table 2 identified 25 genes whose products have been demonstrated to function in cytokine, tumor necrosis factor, T cell receptor signaling, or complement function. STAT, signal transducer and activator of transcription.

rituximab through ADCC, complement-mediated cytotoxicity, or induction of apoptosis.

Prognostic tests allowing the prediction of response to specific therapeutic agents would be of great utility, as they would allow physicians to choose therapy to maximize patient benefit while avoiding unnecessary delay, toxicity, and expense. Our findings suggest that it may be possible, at the time of the initial diagnosis of follicular lymphoma, to predict whether patients will later respond to rituximab treatment for relapsed or refractory disease. The data described here suggest that analysis of gene expression patterns in tumors may allow prediction of the sensitivity of tumors to particular antitumor agents and elucidate the biology underlying resistance to a given therapy. In fact, the possibility of dividing FL into molecularly distinct subtypes may confer additional useful information on differences among these tumors, such as response to therapies other than rituximab. Clearly, our findings require validation on a larger, independent patient cohort. A better understanding of the nature of resistance to rituximab treatment may allow more effective use of this powerful agent. For instance, it is possible that patients who are resistant to standard rituximab dosing may benefit from rituximab administered at a different dose or schedule, or may respond to rituximab in combination with other agents. The ability to prospectively identify patients who are relatively resistant to rituximab would facilitate clinical trials to determine the optimal treatment for such patients.

We thank members of the Brown and Levy laboratories for advice and thoughtful comments on data analysis, and Sandra Horning for thought-provoking discussions. LC microarrays were a much appreciated gift from A. Alizadeh, M. Diehn, and A. Whitney. This work was supported by National Cancer Institute Grants CA85129 (to P.O.B.) and CA77097 (to D.B.), National Institutes of Health Grants CA33399 and CA34233 (to R.L.), and the Howard Hughes Medical Institute. S.P.B. received fellowship support from Amgen and a Public Health Service training grant, and is a Howard Hughes Medical Institute physician postdoctoral fellow. O.A. is an Alfred P. Sloan and U.S. Department of Energy Postdoctoral Fellow in Computational Biology (Grant DE-FG03-99ER62836) and a National Human Genome Research Institute Individual Mentored Scientist Development Awardee in Genomic Research and Analysis (National Institutes of Health Grant K01 HG00038-01). P.O.B. is an Investigator of the Howard Hughes Medical Institute. R.L. is an American Cancer Society Clinical Research Professor.

1. McLaughlin, P., Grillo-Lopez, A. J., Link, B. K., Levy, R., Czuczman, M. S., Williams, M. E., Heyman, M. R., Bence-Bruckler, I., White, C. A., Cabanillas, F., *et al.* (1998) *J. Clin. Oncol.* **16,** 2825–2833.
2. Maloney, D. G., Smith, B. & Rose, A. (2002) *Semin. Oncol.* **29,** 2–9.
3. Clynes, R. A., Towers, T. L., Presta, L. G. & Ravetch, J. V. (2000) *Nat. Med.* **6,** 443–446.
4. Golay, J., Zaffaroni, L., Vaccari, T., Lazzari, M., Borleri, G. M., Bernasconi, S., Tedesco, F., Rambaldi, A. & Introna, M. (2000) *Blood* **95,** 3900–3908.
5. Shan, D., Ledbetter, J. A. & Press, O. W. (2000) *Cancer Immunol. Immunother.* **48,** 673–683.
6. Shan, D., Ledbetter, J. A. & Press, O. W. (1998) *Blood* **91,** 1644–1652.
7. Hofmeister, J. K., Cooney, D. & Coggeshall, K. M. (2000) *Blood Cells Mol. Dis.* **26,** 133–143.
8. Cartron, G., Dacheux, L., Salles, G., Solal-Celigny, P., Bardos, P., Colombat, P. & Watier, H. (2002) *Blood* **99,** 754–758.
9. Perou, C. M., Sorlie, T., Eisen, M. B., van de Rijn, M., Jeffrey, S. S., Rees, C. A., Pollack, J. R., Ross, D. T., Johnsen, H., Akslen, L. A., *et al.* (2000) *Nature* **406,** 747–752.
10. Alizadeh, A. A., Eisen, M. B., Davis, R. E., Ma, C., Lossos, I. S., Rosenwald, A., Boldrick, J. C., Sabet, H., Tran, T., Yu, X., *et al.* (2000) *Nature* **403,** 503–511.
11. Eisen, M. B., Spellman, P. T., Brown, P. O. & Botstein, D. (1998) *Proc. Natl. Acad. Sci. USA* **95,** 14863–14868.
12. Alter, O., Brown, P. O. & Botstein, D. (2000) *Proc. Natl. Acad. Sci. USA* **97,** 10101–10106.
13. Troyanskaya, O., Cantor, M., Sherlock, G., Brown, P., Hastie, T., Tibshirani, R., Botstein, D. & Altman, R. B. (2001) *Bioinformatics* **17,** 520–525.
14. Troyanaskaya, O. G., Garber, M. E., Brown, P. O., Botstein, D. & Altman, R. B. (2002) *Bioinformatics* **18,** 1454–1461.
15. Rosenwald, A., Wright, G., Chan, W. C., Connors, J. M., Campo, E., Fisher, R. I., Gascoyne, R. D., Muller-Hermelink, H. K., Smeland, E. B., Giltnane, J. M., *et al.* (2002) *N. Engl. J. Med.* **346,** 1937–1947.
16. Sorlie, T., Perou, C. M., Tibshirani, R., Aas, T., Geisler, S., Johnsen, H., Hastie, T., Eisen, M. B., van de Rijn, M., Jeffrey, S. S., *et al.* (2001) *Proc. Natl. Acad. Sci. USA* **98,** 10869–10874.
17. Garber, M. E., Troyanskaya, O. G., Schluens, K., Petersen, S., Thaesler, Z., Pacyna-Gengelbach, M., van de Rijn, M., Rosen, G. D., Perou, C. M., Whyte, R. I., *et al.* (2001) *Proc. Natl. Acad. Sci. USA* **98,** 13784–13789.
18. Ghia, P., Granziero, L., Chilosi, M. & Caligaris-Cappio, F. (2002) *Semin. Cancer Biol.*